SOME ARITHMETIC PROPERTIES OF NUMBERS OF THE FORM $\lfloor p^c \rfloor$

WILLIAM D. BANKS, VICTOR Z. GUO, AND IGOR E. SHPARLINSKI

Abstract. Let

$$\mathbb{P}^c = (\lfloor p^c \rfloor)_{p\in\mathbb{P}} \qquad (c > 1,\ c \notin \mathbb{N}),$$

where $\mathbb{P}$ is the set of prime numbers, and $\lfloor\cdot\rfloor$ is the floor function. We show that for every such c there are infinitely many members of $\mathbb{P}^c$ having at most $R(c)$ prime factors, giving explicit estimates for $R(c)$ when c is near one and also when c is large.

1. Introduction

1.1. **Motivation.** *Piatetski-Shapiro sequences* are those sequences of the form

$$\mathbb{N}^c = (\lfloor n^c \rfloor)_{n\in\mathbb{N}} \qquad (c > 1,\ c \notin \mathbb{N}),$$

where $\lfloor t \rfloor$ denotes the integer part of any real number t. Such sequences are named in honor of Piatetski-Shapiro, who showed (cf. [12]) that for any fixed $c \in (1, \frac{12}{11})$ there are infinitely many primes in $\mathbb{N}^c$. The admissible range of c for this result has been extended many times over the years, and currently it is known to hold for all $c \in (1, \frac{243}{205})$ thanks to Rivat and Wu [14].

Many authors have studied arithmetic properties of Piatetski-Shapiro sequences (see Baker *et al* [3] and the references contained therein), and it is natural to ask whether certain properties also hold on special subsequences of the Piatetski-Shapiro sequences. Perhaps the most important of these are the subsequences of the form

$$\mathbb{P}^c = (\lfloor p^c \rfloor)_{p\in\mathbb{P}} \qquad (c > 1,\ c \notin \mathbb{N}),$$

where $\mathbb{P} = \{2, 3, 5, \ldots\}$ is the set of prime numbers; however, up to now very little has been established about the arithmetic structure of $\mathbb{P}^c$ for fixed $c > 1$. Balog [5] has shown that for *almost all* $c > 1$, the counting function

$$\Pi_c(x) = \big|\{\text{prime } p \leqslant x : \lfloor p^c \rfloor \text{ is prime}\}\big|$$

satisfies

$$\limsup_{x\to\infty} \frac{\Pi_c(x)}{x/(c\log^2 x)} \geqslant 1,$$

but this result gives no information for any specific choice of c.

Thanks to the work of Cao and Zhai [7] it is known that the set $\mathbb{P}^c$ contains infinitely many *squarefree* natural numbers provided that c is not too large. More precisely, as a special case of the main result in [7], one knows that for any $c \in (1, \frac{149}{87})$ there exists $\varepsilon > 0$ (depending only on c) such that the estimate

$$\left|\left\{\text{prime } p \leqslant x : \lfloor p^c \rfloor \ \text{ is squarefree}\right\}\right| = \frac{6}{\pi^2} \cdot \pi(x) + O(x^{1-\varepsilon})$$

holds, where $\pi(x)$ denotes the number of primes not exceeding x.

In the present paper, as a step towards better understanding the arithmetic properties of $\mathbb{P}^c$, we consider the related question of whether or not $\mathbb{P}^c$ contains infinitely many *almost primes*.

1.2. **Main results.** For every $R \geqslant 1$, we say that a natural number is an R-*almost prime* if it has at most R prime factors, counted with multiplicity.

We study almost prime values of $\lfloor p^c \rfloor$ in two different regimes in order to demonstrate the underlying ideas: (i) values of c close to one, and (ii) large values of c.

In the first regime, our result is stated in terms of the following set of admissible pairs (R, c_R), $R = 8, \ldots, 19$.

R	c_R	R	c_R	R	c_R
8	1.0521	12	1.1649	16	1.2073
9	1.1056	13	1.1780	17	1.2148
10	1.1308	14	1.1891	18	1.2214
11	1.1494	15	1.1988	19	1.2273

TABLE 1.1. Admissible pairs (R, c_R)

Theorem 1.1. *Let (R, c_R), $R = 8, \ldots, 19$, be a pair from Table 1.1. Then for any fixed $c \in (1, c_R]$ there is a real number $\eta > 0$ such that the lower bound*

$$\left|\left\{\text{prime } p \leqslant x : \lfloor p^c \rfloor \ \text{ is an } R\text{-almost prime}\right\}\right| \geqslant \eta \, \frac{x}{\log^2 x}$$

holds for all sufficiently large x.

In the second regime, we prove the following result.

Theorem 1.2. *For fixed $c \geqslant \frac{11}{5}$ there is a positive integer*

$$R \leqslant \begin{cases} 16c^3 + 179c^2 & \text{if } c \in [\frac{11}{5}, 3), \\ 16c^3 + 88c^2 & \text{if } c \geqslant 3, \end{cases}$$

and a real number $\eta > 0$ such that the lower bound

$$\left|\left\{\text{prime } p \leqslant x : \lfloor p^c \rfloor \text{ is an } R\text{-almost prime}\right\}\right| \geqslant \eta \frac{x}{\log^2 x}$$

holds for all sufficiently large x.

These results are based on bounds of bilinear exponential sums and estimates on the uniformity of distribution of fractional parts $\{p^c d^{-1}\}$. We use the notion of *level of distribution* from sieve theory in a precise form stated in §2.1; see Friedlander and Iwaniec [8] and Greaves [10] We remark that although the ranges of Theorems 1.1 and 1.2 do not overlap, using the same methods and sacrificing on the explicitness of the bounds for R, one can cover the gap as well.

1.3. **Notation.** Throughout the paper, we use the symbols O, $\ll$, $\gg$ and $\asymp$ along with their standard meanings; any constants or functions implied by these symbols may depend on c and (where obvious) on the parameters ε and ν but are absolute otherwise. We use the notation $m \sim M$ as an abbreviation for $M < m \leqslant 2M$.

The letter p always denotes a prime number. As usual, $\mu(\cdot)$ is the Möbius function, and $\Lambda(\cdot)$ is the von Mangoldt function.

We write $\mathbf{e}(t) = \exp(2\pi i t)$ for all $t \in \mathbb{R}$.

2. Proof of Theorem 1.1

2.1. **Preliminaries.** As we have mentioned the following notion plays a crucial rôle in our arguments. We specify it to the form that is suited to our applications; it is based on a result of Greaves [10] that relates level of distribution to R-almost primality. More precisely, we say that an N-element set of integers $\mathcal{A}$ has a *level of distribution D* if for a given multiplicative function $f(d)$ we have

$$\sum_{d \leqslant D} \max_{\gcd(s,d)=1} \left| \left|\{a \in \mathcal{A},\ a \equiv s \bmod d\}\right| - \frac{f(d)}{d} N \right| \leqslant \frac{N}{\log^2 N}.$$

As in [10, pp. 174–175] we define the constants

$$\delta_2 = 0.044560, \qquad \delta_3 = 0.074267, \qquad \delta_4 = 0.103974$$

and

$$\delta_R = 0.124820, \qquad R \geqslant 5.$$

We have the following result, which is [10, Chapter 5, Proposition 1].

Lemma 2.1. *Suppose $\mathcal{A}$ is an N-element set of positive integers with a level of distribution D and degree ρ in the sense that*

$$a < D^\rho \qquad (a \in \mathcal{A})$$

holds with some real number $\rho < R - \delta_R$. Then

$$\left|\{a \in \mathcal{A} : a \text{ is an } R\text{-almost prime}\}\right| \gg_\rho \frac{N}{\log^2 N}\,.$$

Note that we always have $R \geqslant 5$ in what follows.

Using Baker and Pollack [4, Lemma 1] together with Lemma 2.1, it is easily seen that the proof of Theorem 1.1 reduces to showing that, for a fixed pair (R, c_R) as in Table 1.1, for any fixed numbers $c \in (1, c_R]$ and $\vartheta \in (0, 1/R)$ the uniform bound

$$\sum_{1\leqslant h\leqslant H}\sum_{d\sim D}\left|\sum_{n\sim x}\Lambda(n)\mathbf{e}(hd^{-1}n^c)\right| \ll_\vartheta \frac{Dx}{\log^3 x} \tag{2.1}$$

holds with any $D \leqslant x^\vartheta$ and $H = D\log^3 x$. To estimate the triple sums in (2.1) we treat the summation over h with straightforward estimates after estimating the inner sums over d and n. Choosing a sufficiently small $\kappa > 0$ and applying Rivat and Sargos [13, Lemma 2] with

$$\alpha = \max\{1/20, \vartheta + \kappa\} < 1/6,$$

it suffices to show that

$$\sum_{d\sim D} c_d \sum_{m\sim M} a_m \sum_{\ell\sim x/m} b_\ell\, \mathbf{e}(hd^{-1}\ell^c m^c) \ll_{\vartheta,\kappa,\xi} x^{1-\xi} \tag{2.2}$$

with some fixed $\xi > 0$ (depending on ϑ), arbitrary weights c_d, a_m, b_ℓ of size $O(1)$, and in three ranges of M that correspond to two Type I sums and one Type II sum. More precisely, denoting

$$u_0 = x^\alpha$$

these ranges are the following:

(i) Type II sums: $u_0 \ll x/M \ll u_0^2$;
(ii) Type I sums: $u_0^2 \ll x/M \ll x^{1/3}$ with b_ℓ being the characteristic function of an interval;
(iii) Type I sums: $M \ll x^{1/2}u_0^{1/2}$ with b_ℓ being the characteristic function of an interval.

By a standard application of the Fourier analysis (see, e.g., Garaev [9] or Banks *et al* [6]) the hyperbolic region of summation in (2.2) can be replaced with a rectangular region; in other words, it is enough to derive the bound

$$\sum_{d\sim D} c_d \sum_{m\sim M} a_m \sum_{\ell\sim L} b_\ell\, \mathbf{e}(hd^{-1}\ell^c m^c) \ll_{\vartheta,\kappa,\xi} x^{1-\xi} \tag{2.3}$$

for some L and M with $LM \asymp x$ in the following three ranges:

(i) Multilinear Type II sums: $u_0 \ll L \ll u_0^2$;

(ii) Multilinear Type I sums: $u_0^2 \ll L \ll x^{1/3}$ with b_ℓ being the characteristic function of an interval;
(iii) Multilinear Type I sums: $M \ll x^{1/2} u_0^{1/2}$ with b_ℓ being the characteristic function of an interval.

Before proceeding, we record the following technical result which simplifies the exposition below.

Lemma 2.2. *Fix an admissible pair (R, c_R) from Table 1.1. For any fixed numbers $c \in (1, c_R]$ and $\vartheta \in (0, 1/R)$, there is a positive number κ such that if we define*

$$\alpha = \max\{1/20, \vartheta + \kappa\},$$

then all of the following inequalities hold:

(i) $2\vartheta + 2\alpha < c$;
(ii) $c + 5\vartheta + 2\alpha < 2$;
(iii) $365/3 + 32c + 147\vartheta < 174$;
(iv) $8/3 + c + 2\vartheta < 4$;
(v) $2 + c + 4\vartheta < 4$;
(vi) $1 + \vartheta - 2\alpha < 1$;
(vii) $1 + \vartheta/2 - \alpha < 1$;
(viii) $2/3 + \vartheta < 1$;
(ix) $1 - c/2 + 3\vartheta/2 < 1$;
(x) $2\vartheta + (1 + \alpha)/2 < c$;
(xi) $2c + 6\vartheta + \alpha < 3$.

Remark 2.3. These inequalities are listed for convenience only and in some cases are redundant (for instance, (vi) and (vii) are equivalent). The proof of Lemma 2.2 is straightforward.

2.2. **General multilinear sums.** First, we need an adaptation of a result of Baker [1, Theorem 2], which is given here only for the specific exponent pair $(\kappa, \lambda) = (\frac{1}{2}, \frac{1}{2})$. Note that we use D and L instead of M_1 and M_2, respectively in the notation of [1, Theorem 2], and thus we use d and ℓ instead of m_1 and m_2. However, M and m retain the same meaning.

Lemma 2.4. *Let $\alpha_1, \alpha_2, \beta$ be nonzero real numbers such that $\beta < 1$, let h, D, L, M be positive integers, and let g be a real function on the interval $[M, 2M]$ such that*

$$g'(x) \asymp h M^{\beta - j} \qquad (x \sim M).$$

Let

$$S = \sum_{m \sim M} \sum_{d \sim D} \sum_{\ell \sim L} a_m c_{d,\ell} \, \mathbf{e}\left(g(m) d^{\alpha_1} \ell^{\alpha_2}\right)$$

where $a_m, c_{d,\ell}$ are complex numbers with $a_m, c_{d,\ell} \ll 1$. If the number $X = hD^{\alpha_1}L^{\alpha_2}M^{\beta}$ is such that $X \geqslant DL$, then

$$S \ll DLM\left((DL)^{-1/2} + (X/(DLM^2))^{1/6}\right)\log 2DL.$$

Proof. As this is a straightforward variant of [1, Theorem 2] we indicate mainly the changes that are needed in the proof.

Let

$$Q \leqslant DL \tag{2.4}$$

be a natural number to be determined later. Following [1] we see that either (cf. [1, Equation (3.8)])

$$S^2 \ll DLM^2Q\mathcal{L}^2 \tag{2.5}$$

holds with $\mathcal{L} = \log 2DL$ (which corresponds to the value $h = 0$ in [1, Equation (3.6)]), or else we have (cf. [1, Equation (3.9)])

$$S^2 \ll D^2L^2MQ\Delta\mathcal{L}^2\left|\sum_{m\sim M}\mathbf{e}(f(m))\right|, \tag{2.6}$$

where $f(x) = g(x)(d_1^{\alpha_1}\ell_1^{\alpha_2} - d_2^{\alpha_1}\ell_2^{\alpha_2})$ with some quadruple $(d_1, d_2, \ell_1, \ell_2)$ that satisfies

$$d_1, d_2 \sim D, \qquad \ell_1, \ell_2 \sim L, \qquad \Delta - \frac{1}{DL} \leqslant \left|\left(\frac{d_1}{d_2}\right)^{\alpha_1} - \left(\frac{\ell_2}{\ell_1}\right)^{\alpha_2}\right| < 2\Delta$$

where Δ is a number of the form $\Delta = 2^h(DL)^{-1}$ with some fixed integer $h \geqslant 1$, which satisfies the bound

$$\Delta \ll Q^{-1} \tag{2.7}$$

(recall also the condition (2.4)). Note that

$$f'(m) \asymp X\Delta M^{-1} \qquad (x \sim M)$$

as in [1].

Now, if the inequality $X\Delta M^{-1} \leqslant \varepsilon$ holds with for some sufficiently small (but fixed) $\varepsilon > 0$, we can proceed as in Case (i) in the proof of [1, Theorem 2] (making use of [16, Lemma 4.19]) to obtain the bound

$$\sum_{m\sim M}\mathbf{e}(f(m)) \ll X^{-1}\Delta^{-1}M.$$

Since $X \geqslant DL$, upon combining this with (2.6) we again obtain (2.5).

On the other hand, if the inequality $X\Delta M^{-1} > \varepsilon$ holds, then we can proceed as in Case (ii) in the proof of [1, Theorem 2] (with $\kappa = \lambda = \frac{1}{2}$) to derive that

$$\sum_{m\sim M}\mathbf{e}(f(m)) \ll (X\Delta)^{1/2}.$$

Combining this with (2.6) and (2.7) we have

$$S^2 \ll D^2L^2M\mathcal{L}^2(X/Q)^{1/2}. \tag{2.8}$$

Putting (2.5) and (2.8) together, we deduce that

$$S \ll DLM\mathcal{L}\left((Q/(DL))^{1/2} + (X/(M^2Q))^{1/4}\right).$$

The optimal choice for the natural number Q is

$$Q = \left\lceil (D^2L^2X/M^2)^{1/3} \right\rceil.$$

We note that if for the above choice of Q condition (2.4) is not satisfied then $X/M^2 \gg DL$ and the result is trivial. Now, simple calculations lead to the desired bound. □

2.3. **Multilinear sums: Region (i).** In this region, we can apply Lemma 2.4 to bound the sum in (2.3), making the choices $\alpha_1 = -1$, $\alpha_2 = \beta = c$, $c_{d,\ell} = c_d b_\ell$ and $g(x) = hx^c$. Since $LM \asymp x$ and $2\vartheta + 2\alpha < c$ by Lemma 2.2 (i) we see that

$$X = hD^{-1}L^cM^c \geqslant DL \tag{2.9}$$

if x is large, and recalling that $H = D\mathcal{L}^3$ with $\mathcal{L} = \log x$ we also have

$$X \ll HD^{-1}x^c = x^c\mathcal{L}^3;$$

hence, for the sum

$$S = \sum_{d\sim D} c_d \sum_{m\sim M} a_m \sum_{\ell\sim L} b_\ell\, \mathbf{e}(hd^{-1}\ell^c m^c)$$

Lemma 2.4 yields

$$\begin{aligned} S &\ll DLM\mathcal{L}\left((DL)^{-1/2} + (x^c\mathcal{L}^3/(DL))^{1/6}M^{-1/3}\right) \\ &\ll x\left((D/L)^{1/2}\mathcal{L} + (D^5x^c/(LM^2))^{1/6}\mathcal{L}^{3/2}\right). \end{aligned}$$

In Region (i) we have $LM^2 \gg x^2/L \gg x^2u_0^{-2}$, and therefore

$$S \ll x\left((D/L)^{1/2}\mathcal{L} + (D^5x^{c-2}u_0^2)^{1/6}\mathcal{L}^{3/2}\right). \tag{2.10}$$

Recalling our choice of u_0, in Region (i) we have

$$L \geqslant u_0 \geqslant x^{\vartheta+\kappa} \geqslant Dx^\kappa;$$

hence the first term in (2.10) is of size $O(x^{1-\kappa/2})$. For the second term in (2.10), Lemma 2.2 (ii) implies that the inequality

$$5\vartheta + (c-2) + 2\alpha < -\kappa$$

holds with a suitably small κ, hence the second term in (2.10) is of size $O(x^{1-\kappa/2})$ as well.

2.4. **Multilinear sums: Region (ii).** In this region, to estimate the sum

$$S = \sum_{d\sim D} c_d \sum_{m\sim M} a_m \sum_{\ell\sim L} b_\ell\, \mathbf{e}(hd^{-1}\ell^c m^c)$$

in (2.3) we apply a result of Wu [19]. Note that b_ℓ is a characteristic function of an interval. The correspondence between the parameters $(H, M, N, X, \alpha, \beta, \gamma)$ given in [19, Theorem 2] and our parameters is

$$(H, M, N, X, \alpha, \beta, \gamma) \quad \longleftrightarrow \quad (M, D, L, X, c, -1, c)$$

(where $X = hD^{-1}L^cM^c$ as before) and we take $k = 5$ in the statement of [19, Theorem 2]; this gives

$$\begin{aligned} S\mathcal{L}^{-1} \ll \left(X^{32}M^{114}D^{147}L^{137}\right)^{1/174} + (XM^2D^2L^4)^{1/4} + (XM^2D^4L^2)^{1/4} \\ + MD + M(DL)^{1/2} + M^{1/2}DL + X^{-1/2}MDL. \end{aligned}$$

Using the bounds

$$D^{-1}x^c \leqslant X \ll x^c\mathcal{L}^3, \quad LM \asymp x, \quad x^{2\alpha} \ll L \ll x^{1/3} \quad \text{and} \quad D \leqslant x^\vartheta,$$

it follows that

$$\begin{aligned} S\mathcal{L}^{-2} \ll \left(x^{365/3+32c+147\vartheta}\right)^{1/174} + \left(x^{8/3+c+2\vartheta}\right)^{1/4} + \left(x^{2+c+4\vartheta}\right)^{1/4} \\ + x^{1+\vartheta-2\alpha} + x^{1+\vartheta/2-\alpha} + x^{2/3+\vartheta} + x^{1-c/2+3\vartheta/2}. \end{aligned}$$

Taking into account the inequalities of Lemma 2.2 (iii)–(ix) we see that $S = O(x^{1-\kappa})$ if $\kappa > 0$ is small enough.

2.5. **Multilinear sums: Region (iii).** In this region, to estimate the sums in (2.3) we apply a result of Robert and Sargos [15]. Note that b_ℓ is a characteristic function of an interval. The correspondence between the parameters $(H, M, N, X, \alpha, \beta, \gamma)$ given in [15, Theorem 3] and our parameters is

$$(H, M, N, X, \alpha, \beta, \gamma) \quad \longleftrightarrow \quad (D, L, M, X, c, -1, c),$$

where

$$X = hD^{-1}L^cM^c.$$

Applying [15, Theorem 3], for the sum

$$S = \sum_{d\sim D} c_d \sum_{m\sim M} a_m \sum_{\ell\sim L} b_\ell\, \mathbf{e}(hd^{-1}\ell^c m^c)$$

we have the bound

$$S \leqslant (DLM)^{1+o(1)} \left(\left(\frac{X}{DL^2M}\right)^{1/4} + \frac{1}{L^{1/2}} + \frac{1}{X}\right).$$

The third term in this estimate is dominated by the second term since $X \geqslant DL$ (cf. (2.9)), and the second term is dominated by the first term

since $X \geqslant DM$, the latter bound holding in Region (iii) in view of the inequality $c \geqslant 2\vartheta + (1+\alpha)/2$ in Lemma 2.2 (x). Therefore,

$$S \leqslant (DLM)^{1+o(1)} \left(\frac{hD^{-1}L^cM^c}{DL^2M} \right)^{1/4}.$$

Since $h \leqslant H = D\mathcal{L}^3$, $LM \asymp x$, $D \leqslant x^\vartheta$ and $L \gg x^{1/2}u_0^{-1/2}$, we have

$$S \leqslant x^{5/8+c/4+3\vartheta/4+\alpha/8+o(1)}.$$

To prove (2.3) in this case it is enough to show that

$$5/8 + c/4 + 3\vartheta/4 + \alpha/8 < 1,$$

This follows from the inequality

$$2c + 6\vartheta + \alpha < 3,$$

which is given in Lemma 2.2 (xi).

3. Proof of Theorem 1.2

3.1. Preliminaries. Let c be fixed, and put

$$\sigma = \frac{1}{16c^2 + 179c - 1.15c^{-1}} \qquad \text{and} \qquad \beta = 47\sigma. \tag{3.1}$$

For our purposes below, we record that the inequality

$$\frac{c_1(\frac12 - \beta)^3 - (\frac12 - \beta)^4}{(c_1 + \frac12 - \beta)(c_1 + 1 - 2\beta)(2c_1 + \frac12 - \beta)} > \sigma \tag{3.2}$$

holds with $c_1 = c + \sigma$ for all $c \geqslant 2.081$, and the inequalities

$$\frac{\frac{8}{27}c_2 - \frac{16}{81}}{(c_2 + \frac43)(c_2 + 2)(2c_2 + 2)} > 2\sigma \tag{3.3}$$

and

$$\frac{c_2(1 - 2\beta)^3 - (1 - 2\beta)^4}{(c_2 + 2 - 4\beta)(c_2 + 3 - 6\beta)(2c_2 + 3 - 6\beta)} > 2\sigma \tag{3.4}$$

both hold with $c_2 = c - 1 + 3\sigma$ for all $c \geqslant 2.198$.

Suppose that we have the uniform bound

$$\sum_{p \leqslant x} \mathbf{e}(hd^{-1}p^c) \ll x^{1-\sigma} \qquad (d, h \leqslant x^\sigma). \tag{3.5}$$

Let $\mathcal{A}$ be the sieving set given by

$$\mathcal{A} = \big\{ n : n = \lfloor p^c \rfloor \text{ for some prime } p \leqslant x \big\},$$

If (3.5) holds, then (as in the proof of Theorem 1.1) for any fixed $\varepsilon > 0$ we obtain a level of distribution $D = x^{\sigma-\varepsilon}$ for $\mathcal{A}$. Thus, we can apply Lemma 2.1 with $g = c/\sigma + \varepsilon$ (since $a \leqslant x^c$ for all $a \in \mathcal{A}$) and with

$$R \leqslant \frac{c}{\sigma} + 1.15 = 16c^3 + 179c^2, \tag{3.6}$$

which implies the stated result for $c \in [\frac{11}{5}, 3)$.

For $c \geqslant 3$ we replace 179 with 88 in the definition of σ and take $\beta = 20\sigma$ in (3.1), and the estimates (3.4)–(3.6) continue to hold (as well as the bound $\beta < 0.1$; see §3.3 below). Hence, we can also replace 179 with 88 in (3.6) as well.

3.2. **Bounds on some auxiliary sums.** Here, it is convenient to introduce the notations $A \preccurlyeq B$ and $B \succcurlyeq A$, which are equivalents of an inequality of the form $A \leqslant B + O(\mathcal{L}^{-1})$, where $\mathcal{L} = \log N$.

To prove that (3.5) holds, we need the following bound of exponential sums; it is used to establish (3.14) and (3.17) below.

Lemma 3.1. *Let $c, \Theta, \Delta, \varepsilon > 0$ be fixed, and put*

$$k = \lfloor c + \Delta/\Theta \rfloor + 1. \tag{3.7}$$

If $k \geqslant 3$, then the exponential sum

$$S(N) = \sum_{z \sim N^\Theta} \mathbf{e}(z^c N^\Delta)$$

satisfies the bound

$$S(N) \ll N^{\Theta(1-\varrho)}, \tag{3.8}$$

where the implied constant depends only on c and ε, and

$$\varrho = \frac{k-2-\varepsilon}{k(k+1)(2k-1)}. \tag{3.9}$$

Proof. Let $s = k^2 - 1$. Applying the result of Vinogradov [17, Chapter VI, Lemma 7] with the function $F(z) = z^c N^\Delta$ and $n = k$, for any fixed $\varrho \in (0,1)$ we have the bound

$$S(N)^{2s} \ll P^{-2s+\frac{1}{2}k(k+1)} (N^\Theta)^{2s-1+2/k+(k+1)\varrho}\, \mathcal{I} + (N^\Theta)^{2s(1-\varrho)}, \tag{3.10}$$

where

$$\mathcal{I} = \int_0^1 \cdots \int_0^1 \Big| \sum_{z=1}^{P} \mathbf{e}(\alpha_1 z + \cdots + \alpha_k z^k) \Big|^{2s} d\alpha_1 \ldots d\alpha_k$$

and P is the integer given by

$$P = \left\lfloor A_0^{(1-\varrho)/(k+1)} \right\rfloor, \qquad \text{where} \quad A_0 = \left| \frac{(k+1)!}{F^{(k+1)}(N^\Theta)} \right|.$$

Noting that

$$A_0 \asymp N^{\Theta(k+1-c)-\Delta},$$

in order to apply [17, Chapter VI, Lemma 7] it must be the case that

$$\Theta \preccurlyeq \Theta(k+1-c) - \Delta \preccurlyeq \Theta(2+2/k),$$

or in other words,

$$c + \Delta/\Theta \preccurlyeq k \preccurlyeq 1 + 2/k + c + \Delta/\Theta.$$

However, this condition is guaranteed by (3.7).

Applying Wooley [18, Theorem 1.1] with ε/k in place of ε, we see that the integral $\mathcal{I}$ is bounded by

$$\mathcal{I} \ll P^{2s-\frac{1}{2}k(k+1)+\varepsilon/k}. \tag{3.11}$$

Taking into account that

$$P \ll A_0^{1/(k+1)} \ll N^{(\Theta(k+1-c)-\Delta)/(k+1)} \leqslant N^{\Theta},$$

after combining (3.10) and (3.11) we derive the bound

$$S(N)^{2s} \ll (N^{\Theta})^{2s-1+(2+\varepsilon)/k+(k+1)\varrho}\, \mathcal{I} + (N^{\Theta})^{2s(1-\varrho)}.$$

To optimize, we choose ϱ so that

$$2s - 1 + (2+\varepsilon)/k + (k+1)\varrho = 2s(1-\varrho);$$

recalling that $s = k^2 - 1$ this leads to (3.9), and (3.8) follows. □

3.3. **Concluding the proof.** We now turn our attention to (3.5). We use the Heath-Brown decomposition (cf. Heath-Brown [11]) to reduce the problem to that of bounding Type I and Type II sums. In the present situation, to prove (3.5) it suffices to show, for some sufficiently small $\varepsilon > 0$ which depends only on c, that $B = N^{1-\sigma-\varepsilon}$ is an upper bound on all Type I sums

$$S_I(X,Y) = \sum_{x\sim X}\sum_{y\sim Y} a_x\, \mathbf{e}(hd^{-1}x^c y^c) \qquad (Y \gg N^{\frac{1}{2}-\beta}) \tag{3.12}$$

and an upper bound on all Type II sums

$$S_{II}(X,Y) = \sum_{x\sim X}\sum_{y\sim Y} a_x\, b_y\, \mathbf{e}(hd^{-1}x^c y^c) \qquad (N^{2\beta} \ll Y \ll N^{\frac{1}{3}}), \tag{3.13}$$

where $|a_x| \leqslant 1$ and $|b_y| \leqslant 1$, and $XY \asymp N$; we refer the reader to the discussion on [11, pp. 1367-1368]. We specify $\varepsilon > 0$ below.

Let $\mathcal{L} = \log N$ as before. Using van der Corput's inequality with $Q = N^{2\sigma+2\varepsilon}$ and following the proof of Baker [2, Theorem 5], we are lead to the bound [2, Equation (4.18)] with some $q \in [1, Q]$:

$$\begin{aligned} S_{II}(X,Y)^2\mathcal{L}^{-2} &\ll \frac{N^2}{Q} + \frac{N\mathcal{L}q}{Q}\Bigg|\sum_{y\sim Y}\sum_{x\sim X} b_{y+q}\overline{b_y}\, e\big(hd^{-1}x^c((y+q)^c - y^c)\big)\Bigg| \\ &\ll N^{2-2\sigma-2\varepsilon} + N\mathcal{L}\sum_{y\sim Y}\Bigg|\sum_{x\sim X} e\big(hd^{-1}x^c((y+q)^c - y^c)\big)\Bigg|. \end{aligned}$$

For the moment, put $\Theta = (\log X)/\mathcal{L}$, so that $X = N^{\Theta}$. Noting that

$$qY^{c-1}N^{-\sigma} \ll hd^{-1}((y+q)^c - y^c) \ll qY^{c-1}N^{\sigma} \qquad (y \sim Y),$$

and taking into account that

$$Y \asymp N^{1-\Theta} \qquad \text{and} \qquad 1 \leqslant q \leqslant N^{2\sigma+2\varepsilon},$$

we see that in the Type II case it suffices to show that

$$\sum_{z\sim N^{\Theta}} \mathbf{e}(z^c N^{\Delta}) \ll N^{\Theta-2\sigma-3\varepsilon} \tag{3.14}$$

holds uniformly for

$$2/3 \preccurlyeq \Theta \preccurlyeq 1 - 2\beta \tag{3.15}$$

and

$$(1-\Theta)(c-1) - \sigma \preccurlyeq \Delta \preccurlyeq (1-\Theta)(c-1) + 3\sigma + 2\varepsilon, \tag{3.16}$$

where continue to use the notation $A \preccurlyeq B$ from §3.2.

Now put $\Theta = (\log Y)/\mathcal{L}$. Noting that

$$X^c N^{-\sigma} \ll hd^{-1}x^c \ll X^c N^{\sigma} \qquad (x \sim X)$$

and $X \asymp N^{1-\Theta}$, in the Type I case we only need to show that

$$\sum_{z\sim N^{\Theta}} \mathbf{e}(z^c N^{\Delta}) \ll N^{\Theta-\sigma-\varepsilon} \tag{3.17}$$

holds uniformly for

$$1/2 - \beta \preccurlyeq \Theta \leqslant 1 \tag{3.18}$$

and

$$(1-\Theta)c - \sigma \preccurlyeq \Delta \preccurlyeq (1-\Theta)c + \sigma. \tag{3.19}$$

Suppose first that Θ, Δ are such that (3.18) and (3.19) hold, and fix $\varepsilon > 0$. Define k by (3.7) and ϱ by (3.9). Note that $\varrho = f(k)$, where

$$f(t) = \frac{t - 2 - \varepsilon}{t(t+1)(2t-1)}.$$

Since f is decreasing on $[3,\infty)$, and noting that the bounds

$$3 \leqslant k \leqslant c + \Delta/\Theta + 1 \preccurlyeq (c+\sigma)/\Theta + 1$$

hold in view of (3.18) and (3.19), it follows that

$$\Theta\varrho \succcurlyeq f_1(\Theta),$$

where

$$f_1(t) = \frac{c_1 t^3 - (1+\varepsilon)t^4}{(c_1+t)(c_1+2t)(2c_1+t)}$$

with $c_1 = c + \sigma$. Since $c \geqslant 1.6$ and $\varepsilon \leqslant 0.01$ (say), the function f_1 is increasing on $[0,1]$; consequently, as $\Theta \succcurlyeq 1/2 - \beta$ we have

$$\Theta\varrho \succcurlyeq f_1(1/2-\beta) = \frac{c_1(\frac{1}{2}-\beta)^3 - (1+\varepsilon)(\frac{1}{2}-\beta)^4}{(c_1+\frac{1}{2}-\beta)(c_1+1-2\beta)(2c_1+\frac{1}{2}-\beta)}.$$

In view of (3.2) we can choose $\varepsilon > 0$ sufficiently small, depending only on c, such that

$$\Theta\varrho \succcurlyeq \sigma + \varepsilon.$$

Then, using the equation (3.8) of Lemma 3.1, we derive the required bound (3.17) for the Type I sums (3.12).

Next, suppose that Θ, Δ are such that (3.15) and (3.16) hold, and let $\varepsilon > 0$ be chosen as above. We again define k by (3.7) and put $\varrho = f(k)$. Since f is decreasing on $[3,\infty)$, and noting that the bounds

$$3 \leqslant k \leqslant c + \Delta/\Theta + 1 \preccurlyeq (c - 1 + 3\sigma + 2\varepsilon)/\Theta + 2$$

hold in view of (3.15) and (3.16), it follows that

$$\Theta\varrho \succcurlyeq f_2(\Theta),$$

where

$$f_2(t) = \frac{(c_2+2\varepsilon)t^3 - (1+\varepsilon)t^4}{(c_2+2t+2\varepsilon)(c_2+3t+2\varepsilon)(2c_2+3t+4\varepsilon)}$$

with $c_2 = c - 1 + 3\sigma$. Since $c \geqslant \frac{11}{5}$ and $\varepsilon \leqslant 0.01$, one verifies that f_2 attains a unique maximum on $[0,1]$; therefore, as $2/3 \preccurlyeq \Theta \preccurlyeq 1-2\beta$ we have either

$$\Theta\varrho \succcurlyeq f_2(2/3) = \frac{\frac{8}{27}(c_2+2\varepsilon) - \frac{16}{81}(1+\varepsilon)}{(c_2+\frac{4}{3}+2\varepsilon)(c_2+2+2\varepsilon)(2c_2+2+4\varepsilon)}$$

or else

$$\begin{aligned}\Theta\varrho &\succcurlyeq f_2(1-2\beta)\\ &= \frac{(c_2+2\varepsilon)(1-2\beta)^3 - (1+\varepsilon)(1-2\beta)^4}{(c_2+2-4\beta+2\varepsilon)(c_2+3-6\beta+2\varepsilon)(2c_2+3-6\beta+4\varepsilon)}.\end{aligned}$$

In view of the inequalities (3.3) and (3.4), we can take $\varepsilon > 0$ sufficiently small to guarantee that

$$\Theta\varrho \succcurlyeq 2\sigma + 3\varepsilon.$$

Using the equation (3.8) of Lemma 3.1 once again, we derive the required bound (3.14) for the Type II sums (3.13).

Acknowledgements

We thank Roger Baker for his generous help and valuable advice, and for sharing his ideas. In particular, our proofs of the crucial Lemmas 2.4 and 3.1 were originally sketched by Roger Baker. We are also grateful to Xiaodong Cao and Wenguang Zhai for informing us about their paper [7].

During the preparation of this paper, I. E. Shparlinski was supported in part by ARC grants DP130100237 and DP140100118.

References

[1] R. C. Baker, 'The square-free divisor problem,' *Quart J. Math.* **45** (1994), 269–277

[2] R. C. Baker, 'Sums of two relatively prime cubes,' *Acta Arith.* **129** (2007), 103–146.

[3] R. C. Baker, W. D. Banks, J. Brudern, I. E. Shparlinski and A. Weingartner 'Piatetski-Shapiro sequences,' *Acta Arith.* **157** (2013), 37–68.

[4] R. C. Baker and P. Pollack, 'Clusters of primes with square-free translates,' *Preprint*, 2015 (see http://arxiv.org/1505.02744).

[5] A. Balog, 'On a variant of the Pjateckij-Šapiro prime number problem,' in *Groupe de Travail en Théorie Analytique et Élémentaire des Nombres, 1987-1988*, Publ. Math. Orsay, 89-01, Univ. Paris XI, Orsay, 1989, 3–11.

[6] W. D. Banks, J. B. Friedlander, M. Z. Garaev and I. E. Shparlinski, 'Exponential and character sums Mersenne numbers,' *J. Aust. Math. Soc.* **92** (2012), 1–13.

[7] X. Cao, and W. Zhai, 'Distribution of square-free numbers of the form $[n^c]$. II,' *Acta Math. Sinica (Chin. Ser.)* **51** (2008), 1187–1194 (in Chinese).

[8] J. B. Friedlander and H. Iwaniec, *Opera de cribro*, Amer. Math. Soc., Providence, RI, 2010.

[9] M. Z. Garaev, 'An estimate of Kloosterman sums with prime numbers and an application,' *Matem. Zametki* **88** (2010), 365–373 (in Russian).

[10] G. Greaves, *Sieves in Number Theory*. Results in Mathematics and Related Areas (3), **43**. Springer-Verlag, Berlin, 2001.

[11] D. R. Heath-Brown, 'Prime numbers in short intervals and a generalized Vaughan identity,' *Canad. J. Math.* **34** (1982), 1365–1377.

[12] I. I. Piatetski-Shapiro, 'On the distribution of prime numbers in the sequence of the form $\lfloor f(n)\rfloor$,' *Mat. Sb.* **33** (1953), 559–566.

[13] J. Rivat and P. Sargos, 'Nombres premiers de la forme $\lfloor n^c\rfloor$,' *Canad. J. Math.* **53** (2001), 414–433.

[14] J. Rivat and J. Wu, 'Prime numbers of the form $\lfloor n^c \rfloor$,' *Glasg. Math. J.* **43** (2001), 237–254.

[15] O. Robert and P. Sargos, 'Three-dimensional exponential sums with monomials', *J. Reine Angew. Math.* **591** (2006), 1–20.

[16] E. C. Titchmarsh, *The theory of the Riemann zeta-function*. Second edition. The Clarendon Press, Oxford University Press, New York, 1986.

[17] I. M. Vinogradov, *The method of trigonometrical sums in the theory of numbers*, Interscience Publishers, London and New York, 1954.

[18] T. Wooley, 'Vinogradov's mean value theorem via efficient congruencing, II,' *Duke Math. J.* **162** (2013), 673–730.

[19] J. Wu, 'On the primitive circle problem,' *Monatsh. Math.* **135** (2002), no. 1, 69–81.

Department of Mathematics, University of Missouri, Columbia, MO 65211 USA

E-mail address: `bankswd@missouri.edu`

Department of Mathematics, University of Missouri, Columbia, MO 65211 USA

E-mail address: `zgbmf@mail.missouri.edu`

Department of Pure Mathematics, University of New South Wales, Sydney, NSW 2052, Australia

E-mail address: `igor.shparlinski@unsw.edu.au`